AD_____

Award Number:

W81XWH-09-1-0726

TITLE:

SYSTEMS BIOLOGY OF GLUCOCORTICOIDS IN MUSCLE DISEASE

PRINCIPAL INVESTIGATOR:

Zuyi Wang, PhD

CONTRACTING ORGANIZATION:

Children's Research Institute

Washington, DC 20010

REPORT DATE:

October 2010

TYPE OF REPORT:

Annual report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-10-2010	Annual	21 Sept, 2009-20 Sept, 2010

4. TITLE AND SUBTITLE

SYSTEMS BIOLOGY OF GLUCOCORTICOIDS IN MUSCLE DISEASE

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-09-1-0726

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Zuyi Wang, PhD

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Children's Research Institute
Washington, DC 20010

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We have successfully developed and tested the computational genetic network reconstruction algorithms proposed in Aim 1a. Using simulation data, we showed that methods are very effective in reducing noise and estimating regulatory strength and relationship among genes. We believe that our proposed methods are very suitable to the analysis of in vivo transcriptional microarray time series data where various sources of variability in the biological experiments lead to enormous amount of noise. We have been working on identifying gene modules that have significant and distinct change pattern due to the stimulation of glucocorticoids. This step may significantly alleviate the curse of dimensionality. We performed pilot studies on repeated myotoxin injection, BrdU labeling and assessing fiber size distribution in damaged muscle. The results from these experiments are very important and valuable to our ongoing and future study and experiments.

15. SUBJECT TERMS
Duchenne Muscular dystrophy, Glucocorticoids, Systems biology, Drug mechanism

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
U					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	11	**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U			

Table of Contents

	Page
Introduction……..…….	**4**
BODY…….	**4**
Key Research Accomplishments…….…….	**9**
Reportable Outcomes……	**10**
Conclusion……	**10**
References……	**10**
Appendices……	**11**

Annual progress report (2009-2010) for Contract W81XWH-09-1-0726
SYSTEMS BIOLOGY OF GLUCOCORTICOIDS IN MUSCLE DISEASE

Introduction

Duchenne muscular dystrophy (DMD) is the most common and incurable muscular dystrophy of childhood. Muscle regeneration fails with advancing age, leading to considerable fibrosis. Corticosteroid therapy in DMD has been shown to improve muscle strength and function both acutely and over long-time frames. The molecular mechanisms explaining the beneficial effect of glucocorticoids is unknown; and the beneficial response of DMD patients is enigmatic due to the deleterious wasting effects chronic glucocorticoids have on normal muscle. The goal of this proposal is to use computational systems biology approach to model the molecular responses, and to extend and validate the integration of signaling and transcriptional networks using our pre-existing high throughput mouse and rat data sets of transcriptional, proteomic, and PK/PD response to glucocorticoids and human neuromuscular diseases. We also use laboratory approaches to test the hypothesis that the effect of glucocorticoids on reducing inappropriate cross-talk in TGFβ networks improves muscle regeneration. We anticipate that the computational models and experimental results resulted from the proposed research will provide a novel model for efficacy of steroids in DMD.

Body

Specific Aim 1. Provide a multi-scale computational model on the integrated acute proteome (signaling) and transcriptional response regarding metabolic remodeling and TGFβ cascades in muscle induced by glucocorticoids under bolus administration of glucocorticoids. (August 2009 – December 2010)

Progress of Aim 1a. Development of computational molecular network reconstruction algorithms

We have been working on developing the proposed computational network reconstruction algorithms based on microarray gene expression time series. As described in the proposal, our computational network reconstruction approach consists of two components. The first component is gene modulization by wavelet feature extraction, extracting transient features using wavelet transform (Daubechies 1992; Mallat 1999; Percival 2000) from time series signals and modulizing the genes. We will report the preliminary results of gene modulation under Progress of Aim 1b and 1c. The second component is molecular network construction, in which we estimate the structure and parameters of a molecular network to quantify the interaction relationships among the identified genes/gene modules. In the proposal, we planned to use two alternative approaches. Approach A uses DBN to estimate the parameters and conditional probability distributions in Eq. (1) (Kim, Imoto et al. 2003; Perrin, Ralaivola et al. 2003; Beal, Falciani et al. 2005; Zou and Conzen 2005; Fujita, Sato et al. 2007; Heron, Finkenstadt et al. 2007). Our newly proposed Approach B uniquely combines wavelet transform (trend transform) and linear regression for estimating the parameters of the linear regulation model (Eq. (2)).

$$\mathbf{Y}_{t+1} = \mathbf{G}\mathbf{Y}_t + \mathbf{K}\mathbf{U}_t + \mathbf{W}_t, (t = 1,2,...T-1) \quad (1)$$

$$\mathbf{X}_{t+1} = \mathbf{G}\mathbf{X}_t + \mathbf{K}\mathbf{U}_t, \mathbf{Y}_t = \mathbf{C}\mathbf{X}_t + \mathbf{W}_t, (t = 1,2,...T-1) \quad (2)$$

We have first developed and tested Approach B since we have found that Approach B had superior performance than Approach A in the preliminary experiments presented in the proposal. We tested the combined wavelet analysis and linear regression method on simulated data, and obtained promising results. We generated an interaction model (Figure 1b) consisting of six genes (each with a times sequence representing its change in response to an external stimulus) according to the linear Markov model with given parameters and an external stimulus. We simulated different types of the regulatory relationship among the genes and the external stimulus, including positive stimulation and negative inhibition, self-stimulation and self-inhibition, and gene with no interaction with all others (Figure 1b). The thickness of the lines indicates the regulatory strength (regulatory coefficient). Random noises were also added to each signal. Then, we used this method to estimate the regulatory coefficients using the *noisy* signals (i.e., **G**, and **K** in Eq. (2) for wavelet/linear regression). Finally, we used the estimated coefficients to reconstruct the signals, and compared them to the original signals. The closer the reconstructed signals (smooth curves, without noise) are to the

original signals (fluctuating curves, with noise embedded), the better the parameters are estimated. Figure 1a shows that our wavelet/linear regression method accurately reconstructed the signals, and precisely estimated the regulatory strength. The results agree with our expectation that Approach B effectively reduces the impact of large amount of noise and capture change patterns, which makes the estimation of the regulatory strength and relationship more precise.

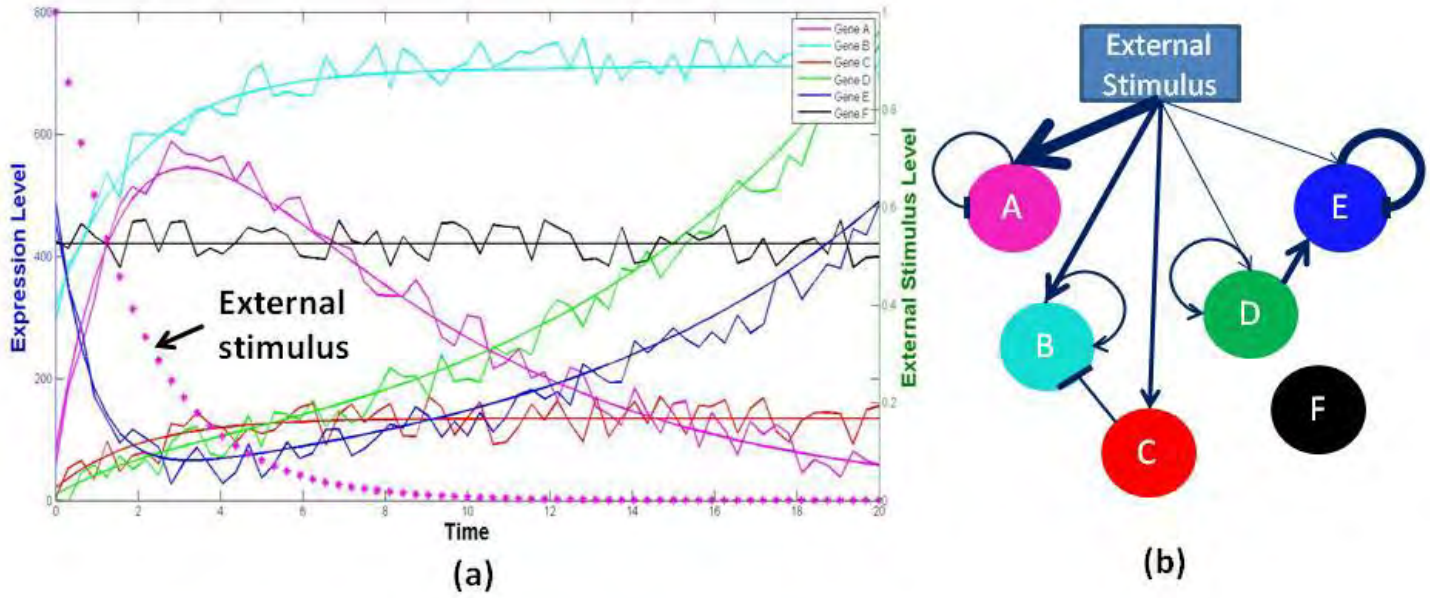


Figure 1. Simulation experiment on genetic network reconstruction using Approach B wavelet/linear regulation model. (a) Simulated gene expression of six genes and an external stimulus. (b) The regulatory relationship among the genes and the external stimulus. The weight of the lines indicates the strength of regulation.

The experiment of Approach A is in progress, and we are seeing more evidence indicating that our Approach B is more effective and efficient. It is likely we will use Approach B instead of Approach A in the future. We are preparing a manuscript to publish this new approach.

<u>Progress of Aim 1b and1c. Applying the computational network reconstruction algorithms developed in Aim 1a to acute transcriptional time series of bolus administration of glucocorticoids.</u>

The work on Aim 1b and 1c is in progress; we report some interim results in the following.

As planned in the proposal, we tried to identify gene modules in the acute transcriptional time series of bolus administration of glucocorticoids. The genes within each gene module have similar change pattern over time, and the change pattern of each gene module is distinct. There is growing evidence that biological networks (metabolic and genetic) function in a multi-scale manner (Jeong, Tombor et al. 2000; Hashimoto, Kim et al. 2004). In the multi-scale context, genes form small groups (or clusters); the constituent genes within each group have close interactions. Some of these clusters form larger meta-clusters that themselves exhibit interactions and this process may continue on several different scales.

To identify the gene modules with distinctive change patterns, we used wavelet analysis to extract features (change patterns) in the time series data. We have been testing the feasibility and effectiveness of different wavelet forms for capturing unique change pattern. Discrete wavelet transform of a discrete time series (signal) consists of trend transform and detail transform; we are using both transform results across multiple levels. Each examines the time series at different locations with different resolutions. The trend series retains the information of the slow varying trend, while the detail series retains the information of fast varying transit (fluctuation). A carefully selected combination of trend and detail components can be used as a “signature” of particular feature of interest. We then use the signature to separate and group the genes into multiple gene modules. Figure 2 shows some interim results of a number of gene modules, each varying in its own specific course after a bolus of glucocorticoids was given. We are refining the results and searching for more gene

modules, and in the meantime we are examining the biological relevance of the genes in each module, We will construct interaction models among these modules and among the genes within each module using the computational methods we developed for Aim 1a.

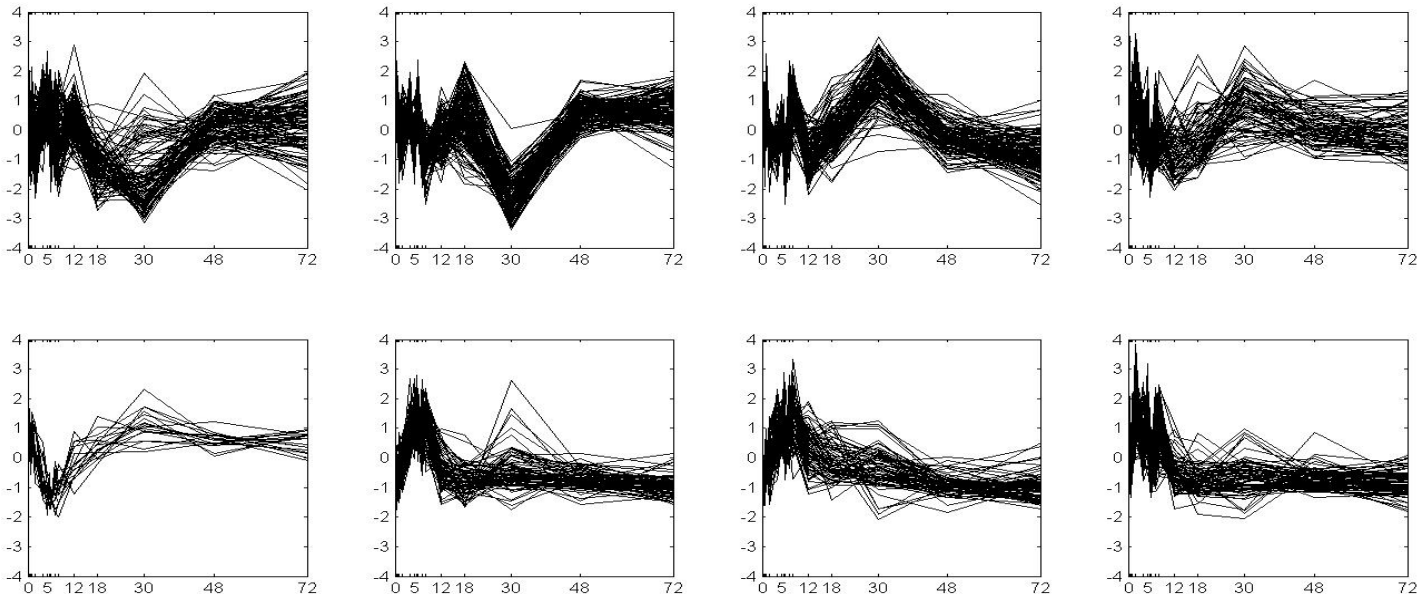


Figure 2. Some gene modules identified in acute transcriptional time series of bolus administration of glucocorticoids using wavelet feature extraction and cluster analysis. The x-axis is time (hours), y-axis is normalized expression level. Each curve of a gene module (in one sub-plot) is the expression change sequence of a

Specific Aim 2. Model the beneficial effects of daily bolus vs. detrimental effects of chronic administration of glucocorticoids. (December 2010 – July 2012)

The work on Aim 2 has not yet started in the first year.

Specific Aim 3. Use a combined modeling and experimental approach to prove the hypothesis that daily dose of glucocorticoids shows efficacy through re-synchronization of the muscle regenerative process. The estimated number of animal used for this Aim is 85. (August 2009 – December 2010)

Progress of Aim 3a. Development of an experimental model of asynchronous muscle regeneration

The work on Aim 3a is ongoing, we report below our findings, problems and difficulties encountered, and potential solutions. It is very important to the success of the study that we ensure that each step of the experiments is carefully planned and all problems are well addressed.

Repeated injection pilot study

We performed a small pilot study to determine if multiple myotoxin injections result in impaired or asynchronous regeneration. Using a total of 10 mice, we bilaterally injected the gastrocnemii of pairs of mice twice, with 1, 2, 4, 5, or 10 days between the first and second injection. The purpose of this spacing was to see if the stage of the regenerating muscle from the first injury resulted in different effects on the second injury. The stages of the first injury when the second injection was place were myoblast activation, myoblast proliferation, myoblast differentiation, myotube fusion, and myotube maturation, respectively.

While performing this study we learned a few methods that will make the future study go more smoothly. We found that dipping the needle in tattoo dye resulted in the surface of the gastrocnemius being covered in dye. The dye on the muscle surface interfered with our ability to place the successive injection accurately because we could not see the original insertion site and only portions of the injection tract through the muscle. Additionally, the tattoo dye on the surface of the muscle dried into a layer of plastic that inhibited proper wound healing. Although we were able to remove the majority of this excess dye during the second surgery, the skin

of some of the animals did not heal together properly. This may have also caused an immune reaction and slowed the muscle regeneration – it is not possible for us to determine the specific effects the excess tattoo dye had. As a result of these two complications, analyzing the tissues has been an extremely slow process. Importantly, we have found that we cannot be sure if the tissue section we are analyzing has been damaged more than once. Based on all of these observations we have come to the decision that future experiments will mix tattoo dye with the myotoxin and be injected into the muscle. While dye dispersion through the tissue may not be to the same extent as the myotoxin, the dye will mark at least a portion of the tissue damaged. Each round of damage will be marked with a different color tattoo dye to enable us to distinguish regions of regeneration. Including the tattoo dye in the injection material will guarantee that the area of the tissue being analyzed has overlapping regions of regeneration because we will be able to see more than one color of dye particles. Notably, we have tested the tattoo dyes and found that they do not interfere with fluorescent immunohistochemical analysis of regeneration. We plan to determine if our other immunohistochemical methods (e.g. hematoxylin and eosin staining) will be affected by the presence of tattoo dye.

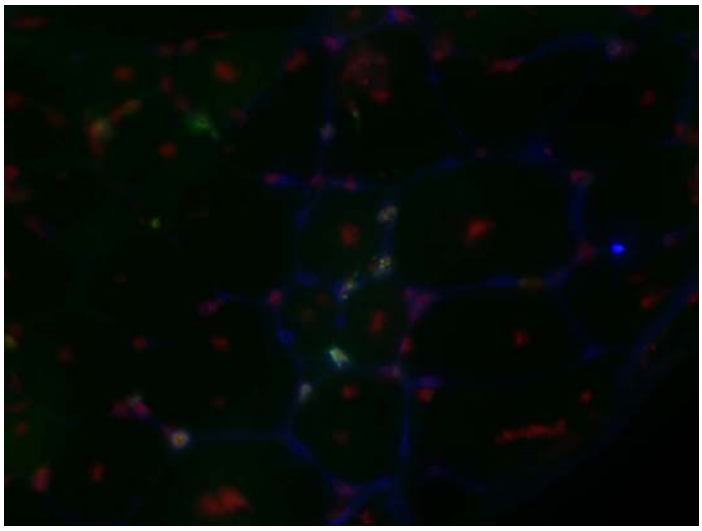
Amendments to the BrdU labeling procedure

While the tissues of this study are currently still being analyzed, we do have some preliminary results. Our initial analysis was to quantify and determine the location of $BrdU^+$ nuclei in regions of the tissue that have regenerated. It should be noted that BrdU was administered three days after the second injection and would only label nuclei that are actively proliferating while the BrdU is present, meaning tissue that regenerated as a result of the first myotoxin injection would be unlabeled. In future experiments we may use other thymidine analogues (e.g. iododeoxyuridine [IUdR]) to label the each round of myogenic regeneration. Additionally, for this experiment we only administered BrdU to the animals twice, as opposed to three times for the previous experiment. It is possible that this hindered our analysis by decreasing the presence of the label in the nuclei and making it harder to determine which nuclei were labeled. The decreased labeling forced us to re-optimize our analysis procedure. Even though we anticipated having these exact problems, we decided to use this lower-level of BrdU procedure because we learned that more than two injections of BrdU in a 24h period can delay tissue regeneration (based on an unpublished study of our colleague, Dr. Marta Fiorotto, in Baylor College of Medicine). We felt that increasing the difficulty of analysis was unfortunate, but it was more important to us to be able to study an unaffected regeneration that would be more scientifically relevant.

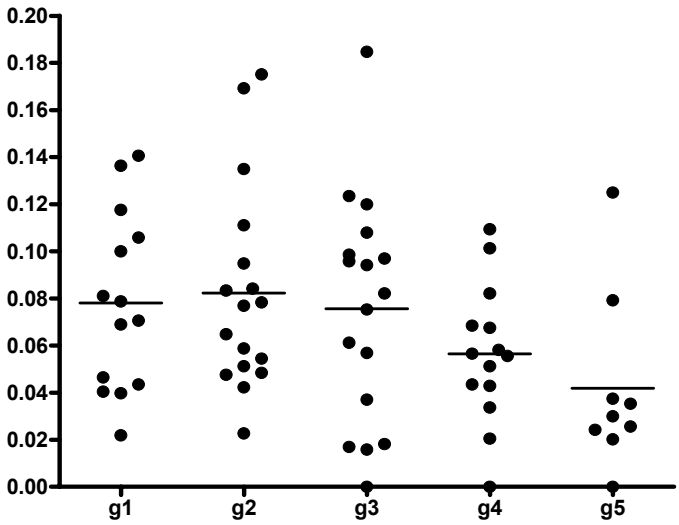
BrdU staining

We have begun our staining of these twice injured gastrocnemii by staining the fiber membranes with anti-laminin, nuclei with propidium iodide (PI), and nuclei that were actively proliferating when we administered BrdU with anti-BrdU (Figure 3). Using laminin to stain myofiber membranes aids in determining where the $BrdU^+$ nuclei are located: centrally within a myofiber, peripherally within the myofiber, or interstitial. At least one example of each can be seen in Figure 3a. We have not completed our analysis of the BrdU staining in these muscles, but do have some preliminary results, such as the ratio of $BrdU^+$ nuclei to total nuclei (Figure 3b). Based on the low ratios of $BrdU^+$ nuclei in our current data set it is likely that we are not yet in the major regions of damage for all of these sections.

In order to determine if we are in the desired area of double injury, we plan to use cathepsin B staining to outline each region of damage. Based on a recent publication from our lab, we plan to use Cathepsin B (CTSB) staining as a marker of inflammation (Baudy, Sali et al. 2010). CTSB is a protease expressed by macrophages. Briefly, we will incubate tissue sections with ProSense - a substrate of CTSB that when cleaved fluoresces in the infra red spectrum – before imaging for CTSB activity. We hope to use this procedure to verify that we are analyzing a region that underwent two times of injury.



(a)



(b)

Figure 3. The results of BrdU labeling of muscle fibers and nuclei with injuries (a) Image of muscle section with BrdU staining. (b) Ratio of $BrdU^+$ nuclei to total nuclei (g1-g5 for different time intervals between two myotoxin injections).

Repeated myotoxin injury induces shift in fiber area.

It is important to note that based on the depth of images taken of the tissue sections we analyze, not all myofibers will have a visible nucleus. For this reason, in order to analyze the pathology of multiply damaged muscle, we measured for the minimum Ferret's diameter of all myofibers. This measurement is commonly used to assess the pathology in dystrophic muscle, which is known to have an abnormal distribution of muscle fiber size – dystrophic muscle being characterized by a higher proportion of smaller fibers, and normal muscle being characterized by a more uniform distribution of fiber size. Additionally, Ferret's diameter is a reliable indicator of cross-sectional area (Briguet, Courdier-Fruh et al. 2004). Although we have begun this analysis, we need to finish before drawing conclusions as generally at least 3000 fiber diameters are measured for this index. Due to the low numbers of fibers measured per study group thus far (Figure 4) we can only describe trends in some of the groups. In order to determine which groups more closely resemble a dystrophic phenotype, we plan to perform these counts in wild-type and mdx mice.

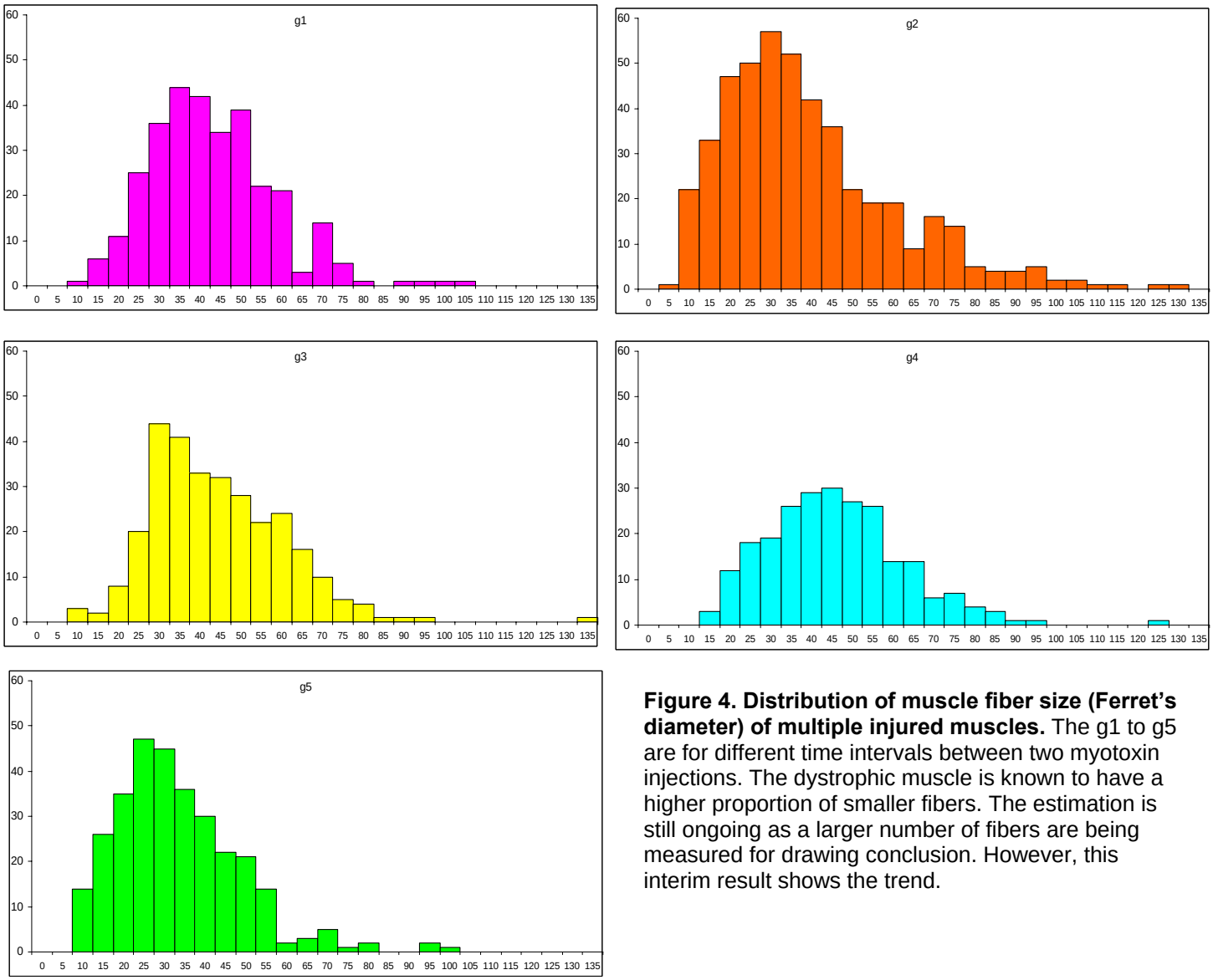


Figure 4. Distribution of muscle fiber size (Ferret's diameter) of multiple injured muscles. The g1 to g5 are for different time intervals between two myotoxin injections. The dystrophic muscle is known to have a higher proportion of smaller fibers. The estimation is still ongoing as a larger number of fibers are being measured for drawing conclusion. However, this interim result shows the trend.

<u>Progress of Aim 3b. Show that daily glucocorticoids result in more successful regeneration in both normal asynchronous regeneration</u>

The work on Aim 3b is ongoing.

Key Research Accomplishments

1. We have successfully developed and tested the computational genetic network reconstruction algorithms proposed in Aim 1a. Using simulation data, we showed that the methods are very effective in reducing noise and estimating regulatory strength and relationship among genes. We believe that our proposed methods are very suitable to the analysis of in vivo transcriptional time series data where various sources of variability in the biological experiments lead to enormous amount of noise. Properly handling noise is decisive to the reliability of genetic networks inferred based on the transcriptional time series data.

2. We have been processing the acute transcriptional time series of bolus administration of glucocorticoids. The main task is to identify gene modules that have significant and distinct change due

to the stimulation of glucocorticoids. We have presented a number of gene modules and their change profiles in this report. This work is ongoing.

3. We performed pilot studies on repeated myotoxin injection and learned a more proper method to make the injection. The new method helps us better distinguish muscle regions of damage of regeneration. The preliminary experimental results from BrdU labeling and staining and muscle fiber size measurement provided valuable information for assessing muscle damage and regeneration.

Reportable Outcomes

1. A new grant awarded by Muscular Dystrophy Association.

   Mechanism of steroid action in DMD, grant # 158914, 1/1/2010 – 12/31/2012.

   Evolved from this award, the MDA grant is a comprehensive project with emphasis on biological validation. The goal is to understand the effects of glucocorticoids on muscle and enable the development of better targeted and more effective therapies for Duchenne muscular dystrophy dynamically. This MDA grant proposal is led by Dr. Eric Hoffman, and it supports a doctoral student's dissertation project.

2. The PI, Dr. Zuyi Wang, and one of the consultants, Dr. Liang Fu, have been preparing a manuscript on the novel computational genetic network reconstruction algorithm proposed in Aim 1a, and plan to submit in 2010.

Conclusion

It is well recognized that reconstructing large scale genetic network is a very complex task. One major obstacle is that the transcriptional microarray data often contain considerable noise that may lead to false estimation of regulatory relationships among genes. Our new method uniquely combines wavelet transform and linear state space model to improve noise resistance and provide accurate estimation of regulatory strength and relationship. Furthermore, we have learned that it is also very important to group genes into modules; this step significantly reduces the problem of curse of dimensionality that is often quite severe in all microarray data. We expect that our algorithms should perform superior when applied to the microarray time series. On the biological experiment part, we learned from our pilot studies on repeated myotoxin injection a more proper method to make the injection. Other experiments on BrdU labeling and fiber size distribution are also very informative for success of future experiments.

References

Baudy, A., A. Sali, et al. (2010). "Non-invasive Optical Imaging of Muscle Pathology in mdx Mice Using Cathepsin Caged Near-Infrared Imaging." Molecular Imaging and Biology: 1-9.

Beal, M. J., F. Falciani, et al. (2005). "A Bayesian approach to reconstructing genetic regulatory networks with hidden factors." Bioinformatics **21**(3): 349-356.

Briguet, A., I. Courdier-Fruh, et al. (2004). "Histological parameters for the quantitative assessment of muscular dystrophy in the mdx-mouse." Neuromuscul Disord **14**(10): 675-82.

Daubechies, I. (1992). Ten lectures on wavelets. Philadelphia, Society for Industrial and Applied Mathematics.

Fujita, A., J. R. Sato, et al. (2007). "Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method." Bioinformatics **23**(13): 1623-1630.

Hashimoto, R. F., S. Kim, et al. (2004). "Growing genetic regulatory networks from seed genes." Bioinformatics **20**(8): 1241-1247.

Heron, E. A., B. Finkenstadt, et al. (2007). "Bayesian inference for dynamic transcriptional regulation; the Hes1 system as a case study." Bioinformatics **23**(19): 2596-2603.

Jeong, H., B. Tombor, et al. (2000). "The large-scale organization of metabolic networks." Nature **407**: 651-4.

Kim, S. Y., S. Imoto, et al. (2003). "Inferring gene networks from time series microarray data using dynamic Bayesian networks." Brief Bioinform **4**(3): 228-235.

Mallat, S. (1999). A Wavelet Tour of Signal Processing, Academic Press.

Percival, D. B. (2000). Wavelet methods for time series analysis. New York, Cambridge University Press.

Perrin, B.-E., L. Ralaivola, et al. (2003). "Gene networks inference using dynamic Bayesian networks." Bioinformatics **19**(suppl_2): ii138-148.

Zou, M. and S. D. Conzen (2005). "A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data." Bioinformatics **21**(1): 71-79.

Appendices

None